
Application of Standard Semantic Web Services and Workflow Technologies in the SIMDAT Pharma Grid

Dr.-Ing. Changtao Qu
C & C Research Laboratories, NEC Europe Ltd.
Sankt Augustin, Germany

W3C Workshop on Frameworks for Semantics in Web Services
Innsbruck, Austria, June 9-10, 2005

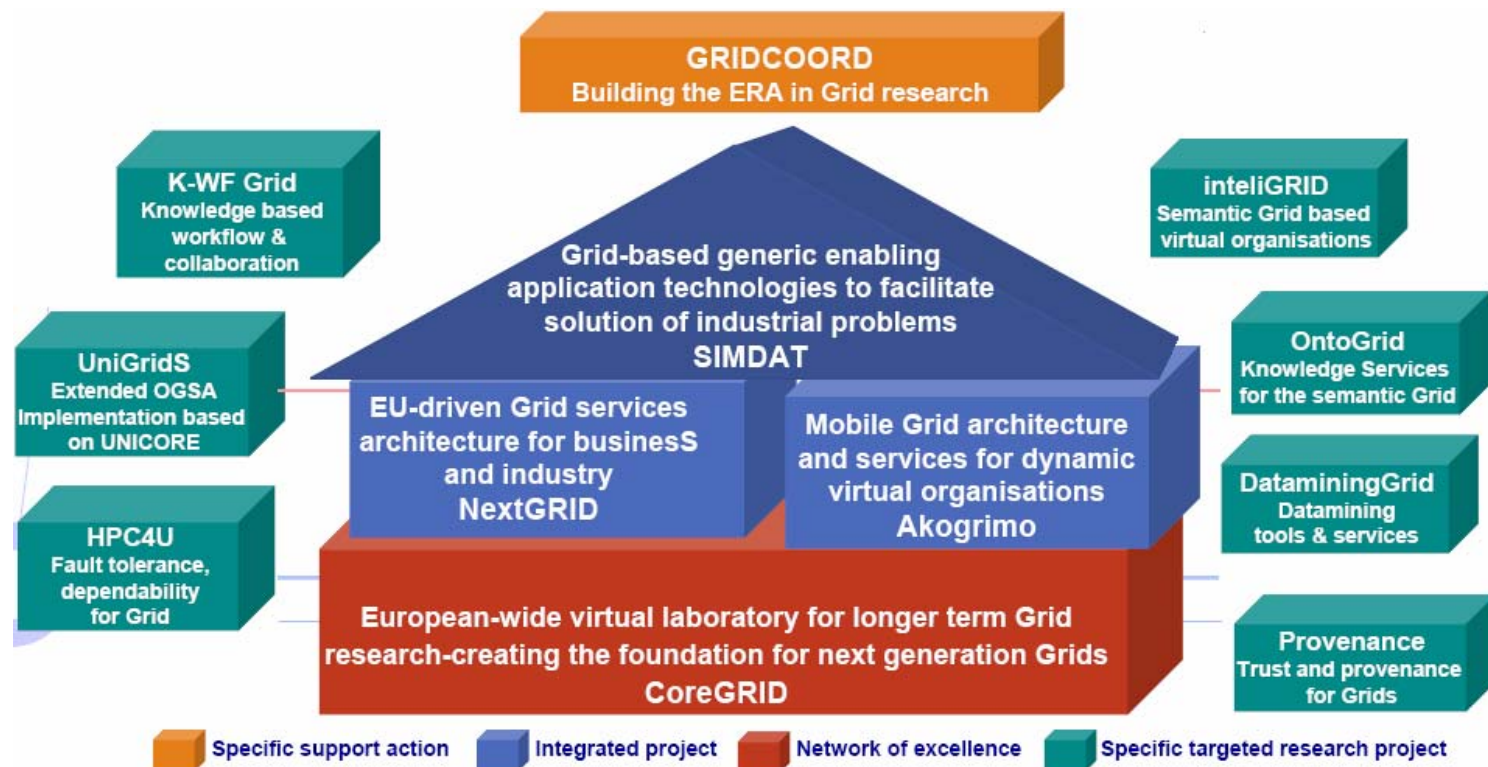


Overview

- SIMDAT Project: Introduction
- SIMDAT Pharma Grid
 - One of the four SIMDAT testbed (i.e. Automotive, Pharma, Aerospace, Meteorology)
- Semantic Broker in the SIMDAT Pharma Grid: Enabling Technologies
 - OWL-DLP/Flogic based ontology
 - OWL-S based service annotation and UDDI based service publication
 - Semantics enabled service matchmaking and composition
 - OWL-S & XScufl based workflow
- Conclusions

SIMDAT Project: Introduction

- SIMDAT: Data Grids for Process and Product Development using Numerical Simulation and Knowledge Discovery
- 4-year EU/IST FP6 project started in Sept. 2004.
- In the EU/IST FP6 Grid project “cluster”, SIMDAT is principally focused on building generic industrial Grid PSEs (Problem Solving Environments).



SIMDAT Project: Introduction

- 4 industrial application sectors, 7 Grid technology development areas

4 industrial application sectors:

- Automotive
- Pharmaceutical
- Aerospace
- Meteorology

7 Grid technology development areas:

- Grid infrastructure
- Distributed Data Access
- VO Administration
- Workflows
- Ontologies
- Analysis Services
- Knowledge Services

- 26 project partners

End Users	Capability Providers	Grid Technology Providers

SIMDAT Pharma Grid: Usage Scenarios

➤ Usage scenario 1: Collaboration of research institutions in EMBNet

- Multiple data and service providers working together to provide genetic researchers with a project working environment.



➤ Usage scenario 2: Pharma/Biotech B2B (GSK)



- Build a platform to distribute data, analysis services and resulting knowledge in an environment that ensures QoS, reliability, authenticity, authorization, and accountability.

➤ Usage scenario 3: Collaboration of global research sites in a Pharma company internal (GSK)



- Multi-site provision and specialized services built on a single architecture to provide proteomics scientists with a specialized resource

SIMDAT Pharma Grid: SRS Service Federation

➤ SRS (Sequence Retrieval System) LION

- Industrial bioinformatics data integration platform, 300+ commercial and academic installations, thousands of users.

The screenshot displays the LION SRS web interface. At the top, there are navigation tabs: 'Quick Searches', 'Predefined Queries', 'Query Builder', and 'Analysis Tools'. Below this, the 'Query Results' section shows a search expression: `{(Bios={UNIPROT}-altwit:"dopamine*")}&{((Bios={UNIPROT}-Organism:"homo*B*sapiens*"|"homo sapiens*"})}`. It reports 'Found 120 entries in Protein Sequence'. A table of results is shown, with columns for 'Accession', 'Description', 'SeqLength', and 'Related Disease entries'. The table lists several entries, including Torsin A precursor, D(2) dopamine receptor, and Dopamine beta-monooxygenase precursor.

Accession	Description	SeqLength	Related Disease entries
UNIPROT:O14656	Torsin A precursor (Torsin family 1 member A) (Dystonia 1 protein).	332	2 entries OMIM:128100 OMIM:605204
UNIPROT:P14416	D(2) dopamine receptor.	443	2 entries
UNIPROT:O15354	Probable G-protein coupled receptor 37 precursor (Endothelin B receptor-like protein-1) (ETBR-EP-1) (Pardir-associated endothelin receptor-like receptor) (PAELR).	613	OMIM:602563
UNIPROT:P32238	Cholecystokinin type A receptor (CCK-A receptor) (CCK-AR).	428	OMIM:119444
UNIPROT:Q99700	Ataxin-2 (Spinocerebellar ataxia type 2 protein) (Tri-nucleotide repeat-containing gene 13 protein).	1312	2 entries
UNIPROT:P28462	D(3) dopamine receptor.	400	OMIM:126451
UNIPROT:P09172	Dopamine beta-monooxygenase precursor (EC 1.3.1.17.1) (Dopamine beta-hydroxylase) (DBH).	608	OMIM:223260

- A typical SRS installation might accommodate 1000+ biological SRS data services and analysis services, implemented as standard Web services.

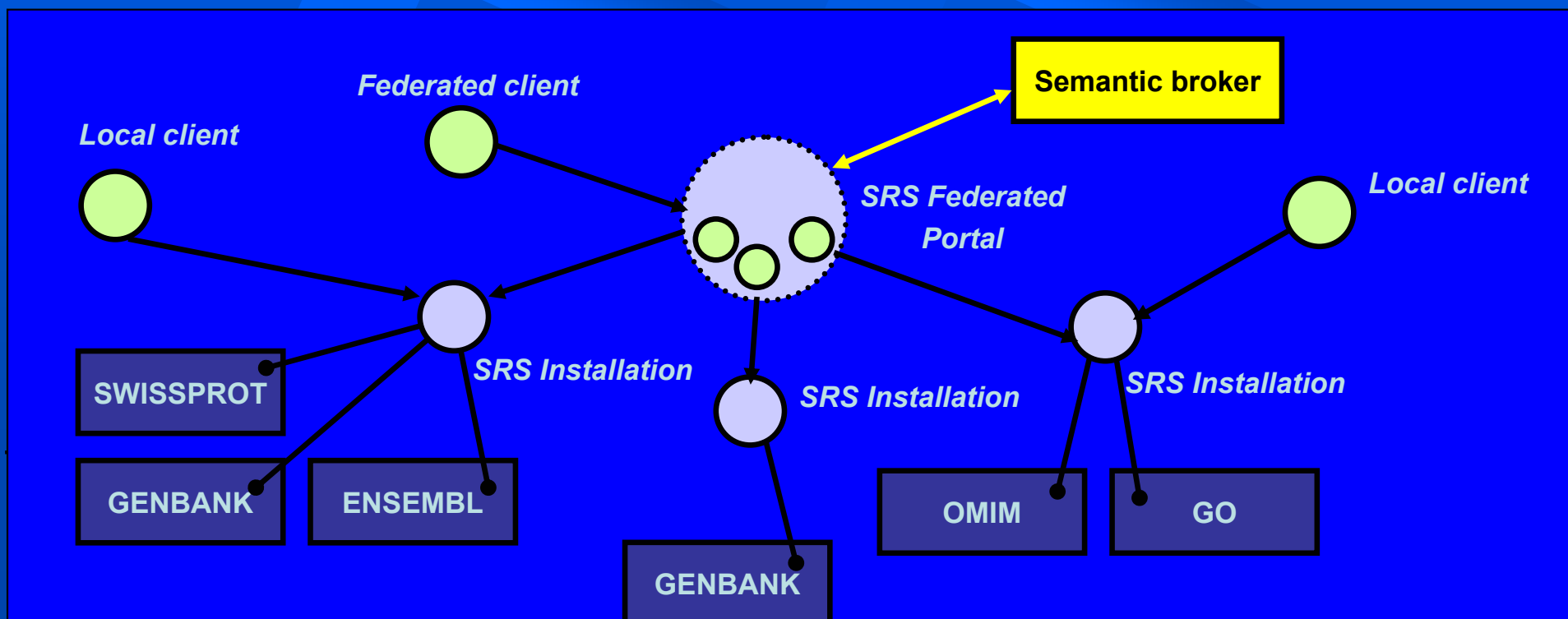
SIMDAT Pharma Grid: SRS Service Federation

➤ Grid-enabled SRS service federation

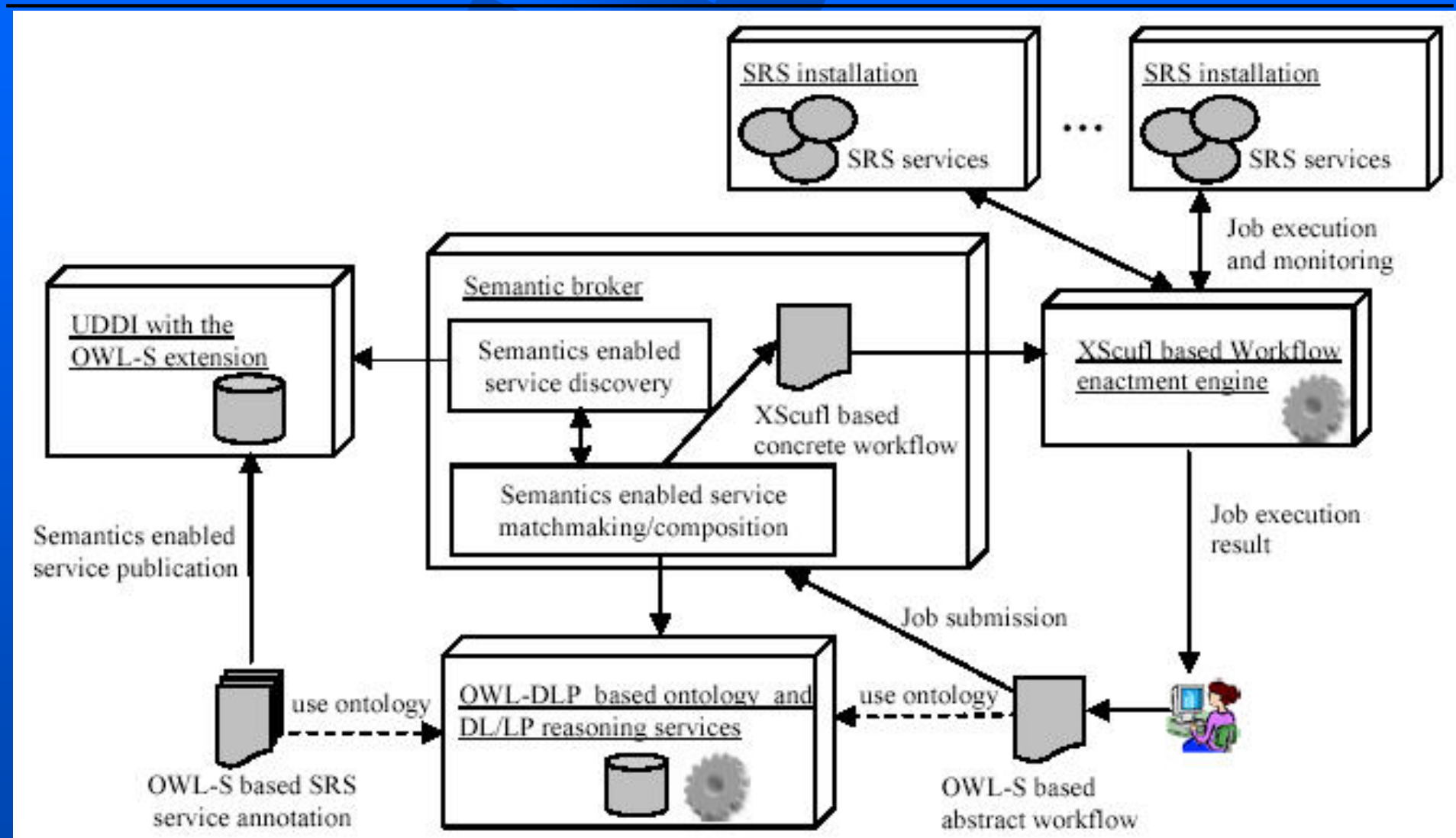
- WS-I/WS-I+ compliant Grid middleware : GRIA (<http://www.gria.org>) **GRIA**
- Grid enabled data access: OGSA-DAI (<http://www.ogsadai.org.uk>) **OGSA-DAI**

➤ Semantic broker

- Intelligently assist Biologists in conducting *in-silico* experiments through automating SRS service discovery, selection, composition and invocation process.





SIMDAT Pharma Grid: Architecture Design



Semantic Broker: Ontology

➤ OWL-DLP/Flogic based ontology

- Develop interoperable, reusable, and maintainable Bioinformatics ontologies.
- Use both DL (Racer/open source) and LP (OntoBroker/commercial) reasoning services
 - 
 - 
- Retain possible interoperability with WSML/WSMO framework.

➤ Distributed and modularized ontology structure

- Cleanly differentiate between SRS application-specific ontology, Bioinformatics domain ontology, and service annotation ontology (OWL-S).
- Reuse existing OWL based ontologies, e.g., vCard, DC, DAML security and privacy.

➤ Development tools:

- Open source: Protégé OWL + Racer
- Commercial: OntoStudio + OntoBroker.

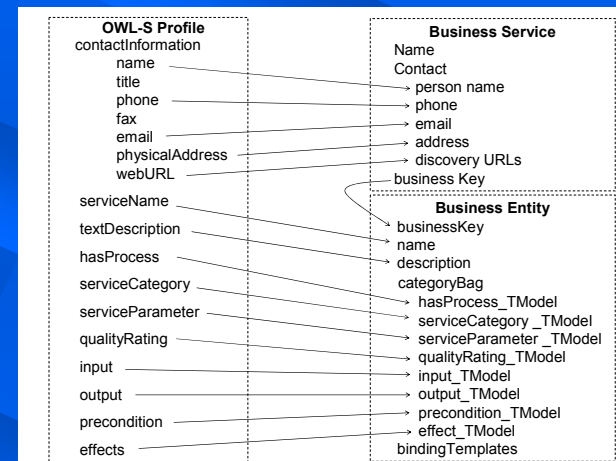
Semantic Broker: Service Annotation & Publication

➤ OWL-S 1.1 based SRS service annotation

- serviceName, textDescription, serviceClassification, serviceCategory (NCBI), serviceParameter (QoS, security requirement, SRS databank), hasInput, hasOutput

➤ UDDI based service publication

- Use the approach proposed by *Paolucci, M., et al., Importing the Semantic Web in UDDI, in Proc. of International Workshop on Web Services, E-Business and Semantic Web, Toronto, Canada, May 2002.*



➤ Development tools:

- Open source: SRI OWL-S editor, CMU OWL-S 1.1 API
- Commercial: OntoStudio.
- Project partner: Fhg SCAI DBAnnotator
- UDDI publication: Apache JUDDI, IBM UDDI4J

Semantic Broker: Service Matchmaking & Composition

➤ Semantics enabled service matchmaking

- Use the approach proposed by *Paolucci, M., et al., Semantic Matching of Web Services Capabilities, in Proc. of International Semantic Web Conference (ISWC), Sardinia, Italy, June 2002.*
- Four level matches: Exact, Plugin, Subsume, Fail.

➤ Service composition

- Not yet addressed at the first stage of the project.
- Might be investigated based on some supporting technologies such as state machines, Petri nets, AI planning, etc.

➤ Development tools:

- Open source: CMU OWL-S 1.1 API, Protégé OWL API 2.0
- Commercial: OntoBroker

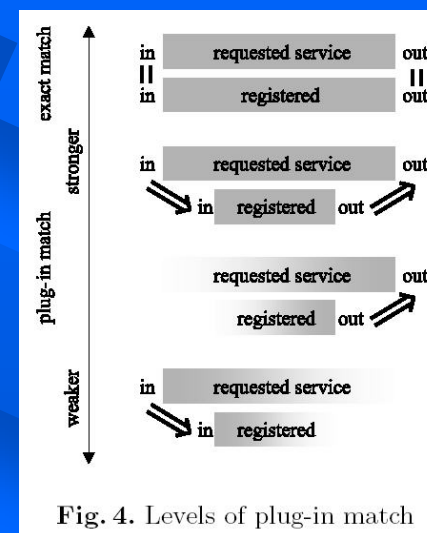




Fig. 4. Levels of plug-in match

Semantic Broker: Workflow

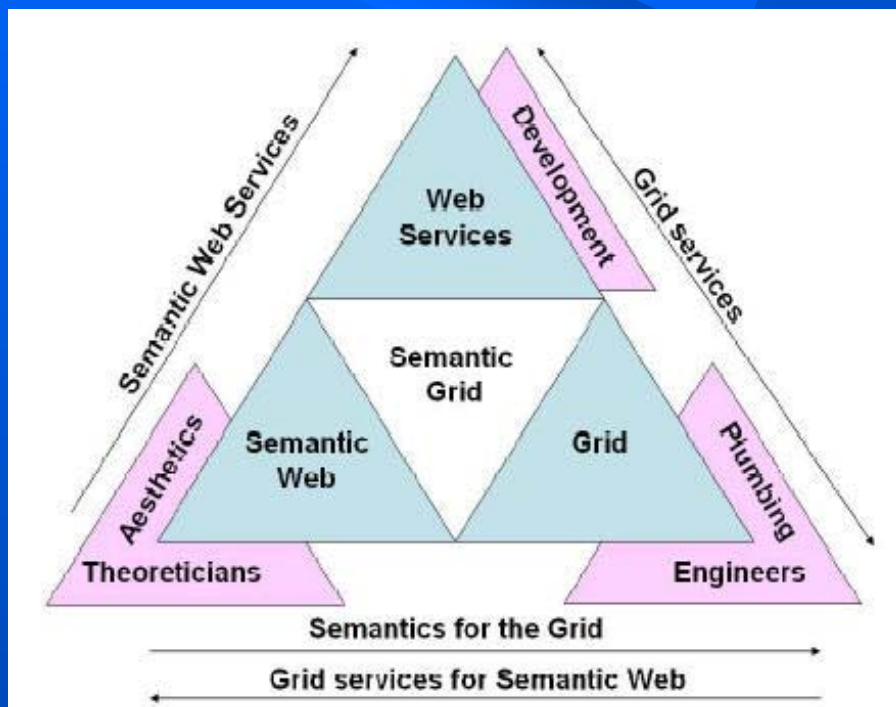
- Workflow description language: OWL-S + XScufl
 - Possibly use OWLS-WS proposed in EU/IST FP6 NextGRID project.
 - Abstract/application workflow Vs. concrete/executable workflow

- Workflow enactment engine: Freefluo/XScufl & InforSense/DPML
 - Implement Grid run-time functionalities such as data staging, service monitoring, QoS negotiation, recovery from faults, etc.

- Development tools:
 - Open source: Taverna (workflow editor), Freefluo 
 - Commercial: InforSense 

Conclusions

- Web, Web services, and Grid; Semantic Web, Semantic Web service, and Semantic Grid.



• C. Goble, and D. De Roure, The Semantic Grid: Myth Busting and Bridge Building , in Proc. of the 16th European Conference on Artificial Intelligence (ECAI-2004), Valencia, Spain, 2004.

- We always have something to learn from each other. We can always benefit from each other.

Thank You for Your Attention



Questions ?

Dr.-Ing. Changtao Qu
 C & C Research Laboratories, NEC Europe Ltd.,
 D-53757, Sankt Augustin, Germany
 Email: qu@ccrl-nece.de Homepage: <http://www.ccrl-nece.de/~qu/>